ORIGINAL ARTICLE

Individualized prediction of illness course at the first psychotic episode: a support vector machine MRI study

J. Mourao-Miranda[1,2], A. A. T. S. Reinders[3,4], V. Rocha-Rego[1], J. Lappin[3], J. Rondina[1], C. Morgan[3], K. D. Morgan[3], P. Fearon[3], P. B. Jones[5], G. A. Doody[6], R. M. Murray[3], S. Kapur[3] and P. Dazzan[3,7]*

[1] *Centre for Neuroimaging Sciences, Institute of Psychiatry, King's College London, UK*
[2] *Centre for Computational Statistics and Machine Learning, UCL, London, UK*
[3] *Department of Psychosis Studies, Institute of Psychiatry, King's College London, UK*
[4] *Department of Neuroscience, University Medical Center Groningen, and BCN Neuroimaging Center, University of Groningen, The Netherlands*
[5] *Department of Psychiatry, University of Cambridge, UK*
[6] *Division of Psychiatry, University of Nottingham, UK*
[7] *NIHR Biomedical Research Centre for Mental Health at the South London and Maudsley NHS Foundation Trust and Institute of Psychiatry, King's College London, UK*

Background. To date, magnetic resonance imaging (MRI) has made little impact on the diagnosis and monitoring of psychoses in individual patients. In this study, we used a support vector machine (SVM) whole-brain classification approach to predict future illness course at the individual level from MRI data obtained at the first psychotic episode.

Method. One hundred patients at their first psychotic episode and 91 healthy controls had an MRI scan. Patients were re-evaluated 6.2 years (S.D. = 2.3) later, and were classified as having a continuous, episodic or intermediate illness course. Twenty-eight subjects with a continuous course were compared with 28 patients with an episodic course and with 28 healthy controls. We trained each SVM classifier independently for the following contrasts: continuous *versus* episodic, continuous *versus* healthy controls, and episodic *versus* healthy controls.

Results. At baseline, patients with a continuous course were already distinguishable, with significance above chance level, from both patients with an episodic course ($p = 0.004$, sensitivity = 71, specificity = 68) and healthy individuals ($p = 0.01$, sensitivity = 71, specificity = 61). Patients with an episodic course could not be distinguished from healthy individuals. When patients with an intermediate outcome were classified according to the discriminating pattern episodic *versus* continuous, 74% of those who did not develop other episodes were classified as episodic, and 65% of those who did develop further episodes were classified as continuous ($p = 0.035$).

Conclusions. We provide preliminary evidence of MRI application in the individualized prediction of future illness course, using a simple and automated SVM pipeline. When replicated and validated in larger groups, this could enable targeted clinical decisions based on imaging data.

Introduction

The outcome of psychosis varies from recovery with minimal symptoms to persistent psychosis with substantial cognitive and functional deficits. Unfortunately, there is no predictor that identifies, at illness onset, which patients will subsequently develop a more severe illness course. Therefore, compared to other areas of medicine, psychiatry lags behind in terms of risk quantification that could guide patients' and clinicians' choices.

In the past decades, brain structure has been investigated using magnetic resonance imaging (MRI) as a potential predictor of outcome in psychosis (van Haren *et al.* 2003, 2008). However, neuroanatomical changes in psychosis are subtle and spatially distributed, although possibly more marked in the advanced illness stages (Ellison-Wright *et al.* 2008; Bloemen *et al.* 2010). As a result, the use of imaging has made little impact in clinical practice on the diagnosis and monitoring of psychoses in individual patients (Matthews *et al.* 2006).

More recently, classification methods have been applied to structural imaging data with encouraging results (Davatzikos *et al.* 2008; Kloppel *et al.* 2008; Vemuri *et al.* 2008; Koutsouleris *et al.* 2009; Karageorgiou *et al.* 2011; Kasparek *et al.* 2011). Among

* Address for correspondence: Dr P. Dazzan, Department of Psychosis Studies, Box 40, Institute of Psychiatry, De Crespigny Park, London SE5 8AF, UK.

(Email: paola.dazzan@kcl.ac.uk)

pattern-based classification algorithms, the support vector machine (SVM) is a pattern recognition technique that identifies multivariate statistical properties of the data that discriminate between groups of subjects, and has shown accuracies up to 86–91% in the correct distinction between patients with established schizophrenia and healthy controls (Fan *et al.* 2008; Sun *et al.* 2009). Although the application of SVM to the classification of patients and healthy controls has validity in research, it has limited clinical utility because the same or even better accuracy can be achieved through clinical interview alone. A more meaningful clinical application of these methods is in helping the identification of predictors of outcome. These could help direct, to more vulnerable individuals, the early implementation of targeted interventions that have been shown to reduce relapse rates, such as optimized pharmacological treatment, assertive case-management or family interventions, resulting in better clinical and functional outcomes (Malla *et al.* 2008; Larsen *et al.* 2011).

In this study, we applied SVM methods to MRI data obtained at the first episode of psychosis, and related them to subsequent illness course type, to provide novel, methodologically sound, proof of concept that MRI can be used at illness onset to predict clinical outcome. Outcome was evaluated on the basis of having achieved symptom control, an important target in the treatment of schizophrenia (Selten *et al.* 2007).

To our knowledge, this is not only the first time that SVM has been used in patients with psychosis to predict clinical outcome but also the first time that it has been applied to MRI scans obtained at the first episode of psychosis, and using a straightforward processing pipeline of promise for application in clinical practice. This application is challenging for classification, as the differences observed at the first episode of psychosis are likely to be much more subtle then those observed in established schizophrenia (Ellison-Wright *et al.* 2008). We predicted that patients with the poorest course (continuously psychotic) could be distinguished both from those with the best course (episodic) and from healthy individuals, with a significant accuracy. In a secondary, exploratory analysis we investigated whether the pattern that best distinguished patients with a continuous course from those with an episodic course could also correctly predict which of the patients with an intermediate course would have developed further illness episodes after the first one.

Method

We recruited a cohort of 100 patients at their first presentation to psychiatric services with a functional psychotic illness (DSM-IV 295–298 psychotic codings; APA, 1994) and 91 healthy individuals with the same socio-economic background. The methods have been extensively described elsewhere (Dazzan *et al.* 2004, 2008). Written informed consent was obtained from all patients, and the study was approved by the local research ethics committee. At first presentation, we interviewed patients using the World Health Organization (WHO) Schedules for Clinical Assessment in Neuropsychiatry (SCAN) and made a diagnosis according to ICD-10 criteria (WHO, 1992*a*, *b*) by consensus in meetings with senior clinicians from the Institute of Psychiatry, in which all clinical information was presented. A total symptomatology score was obtained by summing the SCAN's individual symptom item scores according to the Wing & Sturt (1978) procedure for the Present State Examination (PSE; Wing *et al.* 1974; Wing & Sturt, 1978). This was an appropriate model to adopt as the SCAN incorporates the 10th edition of the PSE. Information on antipsychotic drug dose (in chlorpromazine equivalents) and duration of exposure to antipsychotics was collected during face-to-face interviews, from clinical notes and from interviews with the clinical team.

Evaluation of outcome

Patients were re-evaluated 6.2 (S.D. = 2.3) years later. Information on illness course was obtained using an amended version of the WHO Life Chart (Harrison *et al.* 2001). This measure has been used successfully in previous long-term follow-up studies, and has been shown to be reliable in the assessment of clinical ratings (Susser *et al.* 2000). We used the SCAN (WHO, 1992*b*) criteria to establish absence of psychotic symptoms during the follow-up period, in accordance with WHO and other long-term outcome studies (Jablensky *et al.* 1992). In line with Andreasen *et al.* (2005), we used a 6-month period for establishing remission. The operational criteria used in the study to define a psychotic episode and remission are presented in the Appendix. From the Life Chart, we used illness course type as our primary clinical outcome measure. This was defined as either: continuous (no remission of symptoms of greater than 6 months); episodic (one or more period of remission of at least 6 months, and no episode of psychosis, including the first one, that lasted for 6 months or more); or intermediate (never achieved sustained periods of remission, but also never experienced psychotic symptoms for prolonged periods). We used a conservative approach for the identification of the pattern that best predicted outcome, and only included in the first stage of the SVM analyses patients with either an episodic or a continuous course, thought to best reflect good and
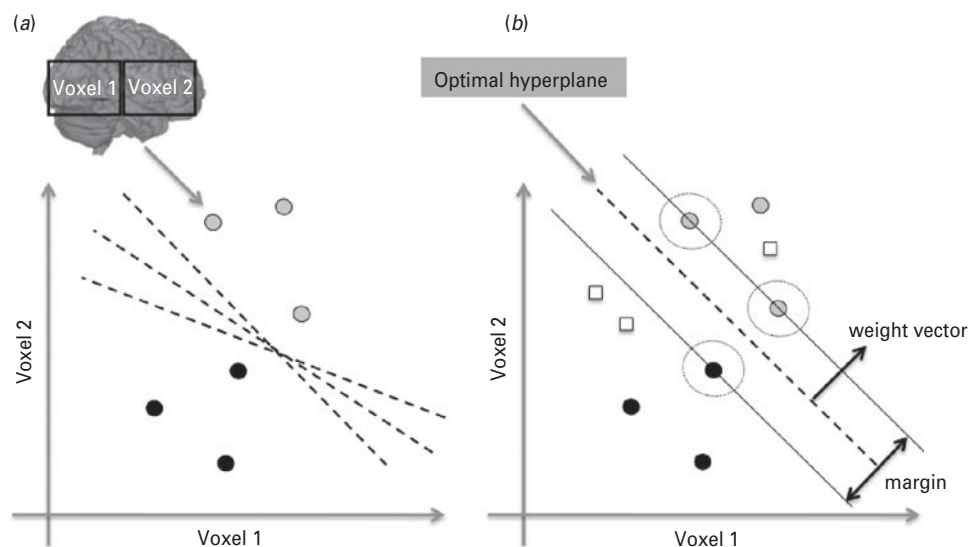
Fig. 1. The support vector machine (SVM) classifier. (*a*) Illustration of a classification problem between two groups (patients *versus* controls) for the simplified case of only two voxels. Each brain image (e.g. gray-matter map) corresponds to a point in the input space and each voxel in the image represents one dimension of this space. The gray circles represent the images of patients and the black circles images of healthy controls. The dashed lines represent hyperplanes or decision boundaries that separate the groups. (*b*) Illustration of the optimal hyperplane determined by the SVM algorithm. The optimal hyperplane (dashed line) is the one with the largest margin of separation between the two classes or groups. The symbols at the margin (circled) are the support vectors. During the training phase the SVM finds the optimal hyperplane or decision boundary. During the test phase the decision boundary can be applied to classify new examples (white squares). The optimal hyperplane is described by a weight vector and an off-set.

poor clinical outcome respectively. In the second stage, we used this pattern to classify the intermediate patients and establish whether the pattern correctly predicted the occurrence of further episodes in this group, during the follow-up period.

Image processing

Images were acquired using a 1.5-T GE scanner (General Electric Medical Systems, USA), at the Maudsley Hospital, London. Tissue maps were produced from T1-weighted coronal images [1.5 mm slice thickness, in-plane resolution 0.9374×0.9374 mm², repetition time (TR) 13.8 ms, inversion time (TI) 450 ms, echo time (TE) 2.8 ms, and flip angle 20°] using SPM5 (Wellcome Department of Cognitive Neurology, London, UK). Modulated plus warped [spatially normalized to Montreal Neurological Institute (MNI) space] gray-matter segments were created with a re-sliced final voxel size of $2 \times 2 \times 2$ mm³, and finally smoothed with an 8-mm isotropic full-width at half-maximum (FWHM) Gaussian kernel.

Pattern classification analysis

Data representation

The modulated, warped and smoothed gray-matter images (probability maps) were used as input to the

SVM classifier. A mask was applied including only gray-matter voxels in common for all subjects (healthy controls and patients with episodic, continuous or intermediate course). To create a common mask, we excluded voxels in the image that had a gray-matter probability value equal to zero for at least one subject.

SVM classification

The SVM is a pattern recognition approach that finds a decision function or boundary that enables classification. It is based on statistical learning theory (Vapnik, 1995) and has emerged as a powerful classification tool. The SVM classifier is trained by providing examples of the form $<\mathbf{x}, c>$ where $\mathbf{x}$ represents a spatial pattern (e.g. gray-matter image) and c is the class label (e.g. $c = +1$ for patients and $c = -1$ for healthy controls). In the context of pattern classification, each brain image corresponds to a point in the input space, and each voxel in the image represents one dimension of this space. A hypothetical example of a classification problem in two-dimensional space is displayed in Fig. 1*a*. The gray circles represent images of patients and the black circles represent images of healthy controls. The dashed lines represent possible separating hyperplanes. During the training phase, the SVM finds the hyperplane or decision function that separates the examples in the input space according to

the group label (e.g. patient *versus* controls). Once the decision function is determined from the training data, it can be used to predict the group membership of a new test example (e.g. white squares, Fig. 1*b*). The optimal hyperplane is described by a weight vector and an off-set. The weight vector is orthogonal to the hyperplane and corresponds to the most discriminating direction between the groups. The weight vector can be plotted as a brain image showing the relative importance of the voxels in discriminating the classes.

In this study we exclusively used a linear kernel SVM to reduce the risk of overfitting the data and to allow direct extraction of the weight vector as an image (the SVM discrimination map). The linear SVM has only one parameter (C) that controls the trade-off between having zero training errors and allowing misclassifications. This was fixed at $C=1$ for all cases (default value). It has been shown previously that the SVM performance for whole-brain classification does not change for a large range of C values and only degrades with very small C values (LaConte *et al.* 2005). This is because there are more dimensions than examples (i.e. more voxels than brain scans). However, in cases where the dimensionality of the data is smaller than the number of examples (e.g. classification based on small regions of interest), it is necessary to use a cross-validation procedure to find the optimal C value. The pattern classification analyses were performed using the PROBID toolbox (www.brainmap.co.uk/probid.htm).

Cross-validation

We trained each classifier independently (continuous *versus* episodic, continuous *versus* healthy controls, episodic *versus* healthy controls). Therefore, each contrast had a different classifier, only used in that contrast. The performance of each classifier was then validated with the commonly used leave-two-out cross-validation approach, which provides a relatively unbiased estimate of the true generalization performance (Ecker *et al.* 2010). In each trial, observations from all but one subject from each group were used to train the classifier. Subsequently, the class assignment of the test subjects was calculated during the test phase. This procedure was repeated S times ($S=$ number of subjects per group), each time leaving out observations from a different subject from each group. The accuracy of the classifier was estimated from the proportion of images correctly classified in both groups, and calculated as the average value of sensitivity and specificity (percentage of test examples from both classes correctly classified). The sensitivity and specificity of the classifier were defined as: sensitivity $=$ TP/(TP$+$FN) and specificity$=$TN/(TN$+$FP), where

TP$=$true positives (proportion of images of group 1 correctly classified); TN$=$true negatives (proportion of images of group 2 correctly classified); FP$=$false positives (proportion of images of group 2 classified as group 1); and FN$=$false negatives (proportion of images of group 1 classified as group 2).

Permutation test

Permutation testing was used to derive a p value for the accuracy of each classifier. Here, we permuted the class labels 1000 times (randomly assigning continuous and episodic labels to the training subjects) and repeated the cross-validation procedure. We then calculated the number of times in which the specificity (percentage of true negative) and sensitivity (percentage of true positive) for the permuted labels were higher than those obtained for the real labels. Dividing this number by 1000 we derived a p value for the classification accuracies.

Applying the SVM classifier to a new sample

In the second stage, the classifier obtained in the continuous *versus* episodic classification was used to classify patients with the intermediate course. In this phase we explored whether these patients were classified as continuous if they had developed further episodes of illness after the first one, or episodic if they had not developed additional episodes.

Discriminating maps (*SVM weight vector*)

The SVM weight vector is a linear combination or weighted average of the support vectors, that is the training examples that are the most difficult to separate. The weight vector is therefore a spatial representation of the decision boundary. Every voxel contributes with a certain weight to the decision boundary or classification function. Given a positive and a negative class (e.g. $+1=$episodic group, $-1=$continuous group), a positive weight for a voxel would indicate that the weighted average in that voxel was higher, for example, for the episodic group, whereas a negative weight would indicate that the weighted average was higher for the continuous group (Mourao-Miranda *et al.* 2005).

The SVM classifiers are multivariate techniques (therefore, they take into account spatial correlations in the data) and because the discrimination is based on the whole brain pattern, rather than on individual regions, all voxels contribute to the classification and no local inferences based on these approaches should be made. We present a list of regions with relatively higher contributions to the decision function or classification. We selected the peaks of the SVM

Table 1. *Sociodemographic and clinical characteristics of the patients included in the analyses*

	Episodic (n=28)	Intermediate (n=32)	Continuous (n=28)	p value
Gender (% males)	50	67	75	N.S.
Age (years)	26.5±25.9	25.9±8.5	28.8±10	N.S.
Education (years)	13.9±2.5	12.2±2.2	12±2.1	0.002
Ethnicity (% White British)	50	41	39	N.S.
Baseline diagnosis (% schizophrenia)[a]	43	66	71	0.07
Follow-up diagnosis (% schizophrenia)[a]	25	75	86	<0.001
Antipsychotic dose (chlorpromazine equivalents, mg)[b]	136±170	197±205	180±212	N.S.
Time on antipsychotics up to MRI (weeks)[c]	4.9±8	10.7±10	6.9±10	0.07
Total symptom severity at baseline	30.5±19	32.1±19	33.2±19	N.S.
Total gray and white-matter volume (liters)	1.14±0.1	1.11±0.1	1.1±0.1	N.S.

MRI, Magnetic resonance imaging; N.S., not significant.

Values given as percentage or mean ± standard deviation.

[a] Change in diagnosis over follow-up does not simply reflect an increase in one diagnostic group and a decrease in another one, but a change in both directions, with approximately 45% of patients changing diagnosis at follow-up.

[b] Information missing for nine subjects.

[c] Information missing for 10 subjects.

weight vector for each classifier, setting the threshold value to 30% of the maximum (absolute) weight value, and estimated the anatomical regions (cluster peaks) that most contributed to the classifier in the discrimination between groups.

Results

From the original cohort, 30 patients developed a continuous course, 35 an episodic course, and 35 an intermediate course. We excluded nine scans because of poor image acquisition and/or motion artifact. In the first stage of the SVM analysis, we compared 28 subjects with a continuous course, 28 with an episodic course, and 28 healthy volunteers [mean age 25.7 (S.D. = 5.6) years; 14 males] of similar age. We then included in the analysis 32 subjects with an intermediate illness course type. Sociodemographic and clinical characteristics of the patient groups are described in Table 1. The mean duration of follow-up was 6.2 (S.D. = 2.3) years. There were more subjects with a baseline diagnosis of schizophrenia spectrum psychosis (schizophrenia, schizophreniform disorder, schizoaffective disorder) in the continuous (71%) and intermediate (66%) groups than in the episodic group (43%, $p = 0.07$). The difference increased further when diagnosis was confirmed at follow-up (continuous 86%, intermediate 75%, episodic 25%, $p < 0.01$).

The first stage of the SVM analysis showed that, at baseline, patients with a future continuous course could already be distinguished, with accuracies above chance level, both from patients with a future episodic course (accuracy = 70%; 71% correctly classified as continuous; $p = 0.004$) and from healthy individuals (accuracy = 67%; 71% correctly classified as continuous; $p = 0.01$) (Table 2). By contrast, patients with an episodic course were not significantly distinguished from healthy individuals (accuracy = 54%).

The anatomical regions with the highest contribution to the discrimination of the continuous group from the episodic and the healthy control groups included the cingulate and parahippocampal gyri, basal ganglia and thalami (Fig. 2, Tables 3 and 4).

In the second and final stage of the analysis we classified patients with an intermediate course according to the discriminating pattern obtained in the episodic *versus* continuous analysis. We found that 78% of those who did not go on to develop other episodes were classified as episodic, and 65% of those who developed further episodes were classified as continuous (sensitivity = 65; specificity = 78; accuracy = 72%; $p = 0.035$). These results indicate that, as expected, training the classifier to discriminate the two 'extreme' groups (continuous *versus* episodic) identifies a pattern that provides information useful to classify the remaining patients.

Discussion

To our knowledge, this is the first study to investigate whether spatially distributed information in brain tissue data, obtained at the first presentation of psychosis, can potentially be used to predict subsequent illness course type at the individual level. The study provides proof of concept that MRI can be used as a tool for the prediction of illness course. Our findings

Table 2. *Results of the support vector machine (SVM) classification*

Comparison	Sensitivity	Specificity	Accuracy	p
Continuous *versus* episodic	71	68	70	0.004
Continuous *versus* healthy individuals	71	61	67	0.01
Episodic *versus* healthy individuals	64	43	54	0.3

The results are give for the SVM classification in two classes: ratio of true positive (sensitivity), true negative (specificity), accuracy and statistical probability. The first column shows the groups of subjects considered in each classification. The second column shows the percentage of subjects in the first group correctly classified as pertaining to it (sensitivity). The third column shows the percentage of subjects in the second group correctly classified as non-pertaining to the first group (specificity). The fourth column shows the accuracy (arithmetic mean between sensitivity and specificity). The last column shows the statistical probability that the result has been obtained by chance. It was obtained after 1000 permutations within the subjects. The number of subjects considered in each classification was 56 (28 in each group).
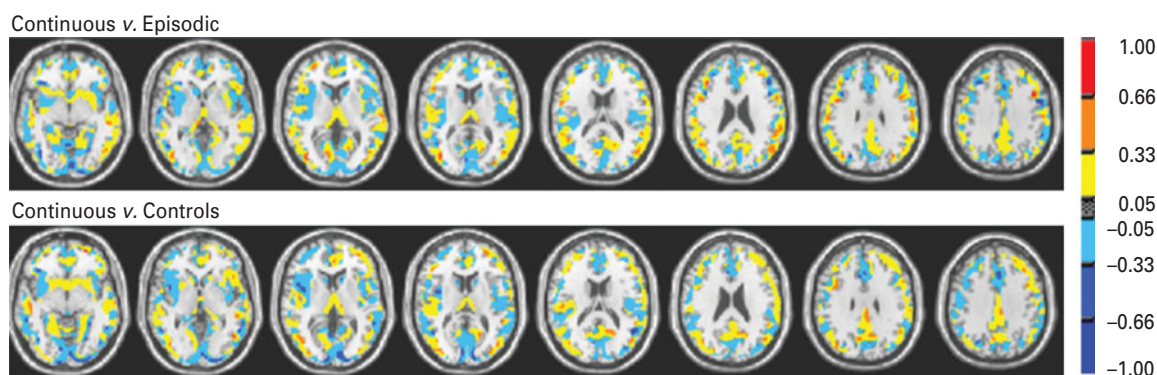

Fig. 2. Discrimination map or support vector machine (SVM) weight vector: continuous *versus* episodic course (*top*), continuous course *versus* healthy individuals (*bottom*). The colours represent the weight of each voxel in the classification function (the red scale represents positive weights and the blue scale represents negative weights). The SVM weight vector is a linear combination or weighted average of the support vectors, that is the training examples that are most difficult to separate and define the decision boundary. The weight vector is therefore a spatial representation of the decision boundary. Every voxel contributes with a certain weight to the decision boundary or classification function. Given a positive and a negative class (e.g. +1=episodic group; −1=continuous group), a positive weight for a voxel means the weighted average in that voxel was higher for the episodic group, and a negative weight means the weighted average was higher for the continuous group. Because the classifier is multivariate by nature, the combination of all voxels as a whole is identified as a global spatial pattern by which the groups differ (the discriminating pattern). Therefore, the discrimination map should not be interpreted as a standard statistical parametric map resulting from a mass-univariate statistical test to find group differences, and no local inferences should be made based on the SVM weights.

suggest that at the time of first presentation, patients who will subsequently develop a continuous illness course may be already distinguished, above chance level, from those who will develop an episodic course, using a classifier trained only on patient groups. By contrast, patients with a future episodic course cannot be distinguished from healthy individuals. Finally, our findings provide preliminary evidence that, even for patients with a 'less extreme' (or intermediate) type of outcome, it is possible to predict, to a certain extent, who will develop further episodes and who will be less likely to do so.

Although significant, the accuracy with which we have correctly classified patients is lower than that reported in other studies that have used SVM (Fan *et al.* 2008; Sun *et al.* 2009). Of note, these studies have applied SVM to patients with an established illness, rather than at first episode, and have used SVM to define subject status (patient *versus* control) rather than future illness course. It is possible that, as more marked and distributed brain changes occur with illness progression (Ellison-Wright *et al.* 2008), the classification and distinction between patients and healthy individuals becomes more accurate in later illness stages. Interestingly, a study that applied SVM to the early detection of different at-risk mental states for subthreshold psychotic symptoms reported accuracies lower than those reported in patients with established

Table 3. *List of the most discriminating regions (cluster peaks) for the classifier episodic versus continuous*

Regions	x	y	z	Cluster volume (mm³)	Peak value
Positive weights					
Right caudate nucleus	18.8	4.9	16.9	770	17.045
Left caudate nucleus	−19.8	13.8	12	88	14.899
Left putamen	−20.8	1.9	13.9	140	15.235
Right putamen	29.7	−8	3.1	130	8.5717
Substantia nigra	9.9	−15.9	−12.8	426	11.502
Right globus pallidus	25.7	−13	−3.9	98	9.9449
Right anterior cingulate cortex (BA 25)	5.9	11.8	−10.8	324	10.65
Right parahippocampal gyrus (BA 30, BA 27)	15.8	−41.7	6	330	13.916
	20.8	−34.8	−3.9	105	9.8681
Left parahippocampal gyrus (BA 30, BA 28, BA 34)	−19.8	−36.7	5	173	16.723
	−16.8	−11	−11.8	86	8.2407
	−23.7	0.9	−10.8	56	10.665
Negative weights					
Left caudate nucleus	−16.8	−23.9	27.8	491	−12.158
Right caudate nucleus	18.8	1.9	22.8	186	−19.143
Left putamen	−27.7	−18.9	−5.8	81	−11.904
Right putamen	29.7	−14	−7.8	85	−10.77
Right globus pallidus	11.9	−7	−7.8	85	−9.2672
Right anterior cingulate gyrus (BA 24)	9.9	−1.1	33.7	329	−15.372
Left anterior cingulate gyrus (BA 24)	0	−7	22.8	91	−11.821
Left posterior cingulate gyrus (BA 29)	−4	−41.7	17.9	53	−9.1322
Right posterior cingulate gyrus (BA 23)	2	−23.9	26.8	181	−10.48
Left parahippocampal gyrus (BA 27, BA 35)	−19.8	−33.8	−3.9	97	−13.182
	−22.7	−28.8	−11.8	137	−9.364
Left amygdala	−26.7	−9	−11.8	114	−10.084
Right amygdala	27.7	−7	−11.8	279	−12.003
Hypothalamus	−6.9	−3.1	−10.8	69	−11.211
Left cerebellum	−15.8	−41.7	−16.7	188	−9.8005
Right cerebellum	5.9	−41.7	−10.8	351	−11.505

x, y, z are Talairach coordinates of the cluster peaks selected using 3Dclust in AFNI (http://afni.nimh.nih.gov/afni). The regions were estimated using the software Talairach Client (www.talairach.org/).

schizophrenia (Koutsouleris *et al*. 2009). Consistent with our finding, a study that used a different classification approach to predict 1-year outcome in first-episode psychosis patients also reported an accuracy similar to the one we observed (Kasparek *et al*. 2011). Nevertheless, the better classification we found in individuals destined to develop a continuous illness course suggests that significant brain changes have already occurred by the time of first contact in this subgroup.

Although, with these methods, it is not possible to make local inferences on the discriminating regions, it is noteworthy that the anatomical regions with the highest contribution to the discrimination between groups include the cingulate and parahippocampal gyri, the basal ganglia and the thalami, which have been consistently reported as important in the pathophysiology of schizophrenia (Ellison-Wright *et al*. 2008).

Our study should be taken as preliminary, but it proposes a promising approach for the future translation of imaging into patient benefit. Although our approach requires replication and validation in a larger sample, it provides initial evidence of an easy and accessible methodology that could potentially enable rapid clinical decisions. We believe that future development should consider that, although a better classification accuracy could potentially be achieved with feature selection approaches preceding the SVM classifier, this would involve applying cross-validation pipelines that are time-consuming, calculation intensive and/or methodologically challenging. Such approaches can be useful in the research setting but might not enable a rapid and user-friendly assessment

Table 4. *List of the most discriminating regions (cluster peaks) for the classifier continuous versus healthy individuals*

Regions	x	y	z	Cluster volume (mm³)	Peak value
Positive weights					
Right caudate nucleus	19.8	4.9	19.9	660	16.11
Left pulvinar	−19.8	−35.8	2.1	581	12.659
Right pulvinar	10.9	−28.8	10	552	15.777
Left putamen	−18.8	2.9	13.9	210	14.144
Right putamen	22.7	11.8	−2.9	121	8.363
Left lateral globus pallidus	−24.7	−9	−2.9	164	12.189
Right thalamus	2	−13	15.9	432	9.8022
	23.7	−20.9	7	145	15.65
Left red nucleus	−3	−22.9	−4.8	310	10.292
Right parahippocampal gyrus (BA30, BA28)	16.8	−40.7	5	282	14.279
	−15.8	−6	−10.8	161	10.278
Left parahippocampal gyrus (BA30)	−15.8	−35.8	−6.8	61	8.3901
Right anterior cingulate gyrus (BA24)	11.9	27.7	−2.9	229	16.131
Right posterior cingulate gyrus (BA 23)	5.9	−20.9	22.8	78	14.808
Left cerebellum	−5.9	−35.8	2.1	57	9.5645
Negative weights					
Left caudate nucleus	−4	12.8	13	722	−11.528
Right caudate nucleus	18.8	−23.9	22.8	153	−9.527
Left putamen	−19.8	14.8	−2.9	500	−10.024
Right putamen	27.7	−22.9	1.1	243	−13.137
Left pulvinar	−13.8	−34.8	16.9	99	−13.191
Right pulvinar	21.8	−27.8	9	124	−21.096
Right globus pallidus	20.8	−10	8	64	−8.4599
Hypothalamus	−6.9	−3.1	−9.8	189	−9.0975
Left anterior cingulate gyrus (BA24)	−1	−3.1	27.8	107	−11.888
Right anterior cingulate gyrus (BA24)	12.9	−4	32.7	154	−15.651
Parahippocampal gyrus (BA 25, BA28)	−21.8	−16.9	−9.8	517	−15.527
	−15.8	−22.9	−21.7	55	−7.2935
Cerebellum	6.9	−42.7	−10.8	1795	−13.963

x, y, z are Talairach coordinates of the cluster peaks selected using 3Dclust in AFNI (http://afni.nimh.nih.gov/afni). The regions were estimated using the software Talairach Client (www.talairach.org/).

of patients' prognosis. Furthermore, there is at present no conclusive evidence that feature selection increases the accuracy for MRI-based diagnosis, while substantially increasing computation time without improving classification performance (Cuingnet *et al.* 2011). Therefore, we have preferred here a whole brain approach, which has stronger potential for future applicability in the clinical setting. Once the SVM classifier is trained and a decision function is generated, a new example could be classified in minutes. The pipeline we used includes pre-processing of the structural images using standard neuroimaging software (SPM), extraction of a feature vector containing gray-matter values, and application of the SVM decision function to the feature vector. Moreover, the greater the complexity of the method (i.e. the number of steps in the pipeline), the greater the danger of overfitting the data. By contrast, maintaining a simple

pipeline improves the translational potential of the method, as such a technique will be more likely to be incorporated into routine clinical investigations, because a structural MRI can be obtained in 10 minutes.

In psychiatry, clinicians do not have algorithms to quantify the risk of, for example, occurrence of further episodes, which could help in guiding clinical decisions. Such algorithms are of common use in other areas of medicine, such as heart medicine and oncology. Our results offer preliminary evidence that, once further developed and validated, an SVM approach could in the future be used by a psychiatrist to say, for example, that on the basis of the MRI scan, a patient may have 88% probability of having more illness episodes (positive predictive value: in this example, it is calculated as the proportion of intermediate subjects who went on to experience repeated episodes and who were classified as continuous by the

SVM classifier), thus providing more information on which to base therapeutic decisions. This would imply applying risk quantification in psychiatry, something that has not yet been achieved. Furthermore, the quantification could improve even further and be refined if other factors related to outcome, such as duration of untreated psychosis, were taken into account in a decision tree.

As mentioned earlier, our study is preliminary and the findings should be considered with caution. An important limitation is the relatively small sample size. In addition, future work should consider validating the accuracy of our classifier with an independent larger sample of patients classified according to the same illness course criteria. Finally, as this study used data from a single site, it remains unclear to what extent differences in acquisition protocol or scanners affect the accuracy of the classifier. Nevertheless, it is encouraging that the classifier trained in using the two extremes groups of patients (i.e. episodic *versus* continuous) could be successfully applied to classify the intermediate group of patients. Future studies with multi-site images will allow an investigation of the extendibility of the approach to images from different sites.

Ultimately, once developed, these methods could allow a patient to be assigned to targeted assertive case-management at first presentation to services, with optimized pharmacological treatment, cognitive and family interventions, which have been shown to improve treatment adherence and reduce relapse rates (Malla *et al.* 2008; Eack *et al.* 2010), eventually improving outcome. At the same time, those patients most likely to have a good remitting illness after their first episode could avoid long-term exposure to antipsychotic medication. We propose that, with further development and validation on larger datasets, a simple and automated SVM pipeline offers a promising approach to help rapid and early clinical decisions based on imaging data, which will eventually benefit patient care and reduce health-care costs.

Appendix

The operational definition of a 'psychotic episode'

A psychotic episode is a period of symptomatology in which the presence of at least one of the following symptoms can be ascertained:

Group A

1. Hallucinations or pseudo-hallucinations in any modality.
2. Delusions.

3. Marked thought and speech disorder (e.g. incoherence, irrelevance, thought blocking, neologisms, incomprehensibility of speech) other than simple retardation or acceleration.
4. Marked psychomotor disorder (e.g. negativism, mutism or stupor, catatonic excitement, constrained attitudes or unnatural postures maintained for long periods) other than simple retardation or acceleration.
5. Emergence or marked exacerbation of bizarre and grossly inappropriate behavior (e.g. talking or giggling to self, acts incomprehensible to others, loss of social constraints, etc.).

A psychotic episode may be considered as present also in the absence of the manifest symptoms listed in Group A if at least two of the following behaviors have emerged or become markedly exacerbated:

Group B

1. Marked reduction or loss of interest, initiative and drive, leading to serious deterioration of the performance of usual activities and tasks.
2. Emergence or marked exacerbation of social withdrawal (active avoidance of communication with other people).
3. Severe excitement, purposeless destructiveness or aggression.
4. Episodic or persistent states of overwhelming fear or severe anxiety.
5. Gross and persistent self-neglect.

The operational definition of a 'remission'

A remission is a state following a psychotic episode, in which none of the symptoms listed as characteristics of a psychotic episode are present. During a remission a patient may exhibit a variety of non-psychotic symptoms (e.g. depressed mood, neurotic manifestations) or some of the so-called negative symptoms, or be entirely symptom free (incomplete or complete remission). A rating of remission (and also a rating of a psychotic episode) should be based only on the ascertainable absence (or presence) of psychotic symptoms and not on whether or not the patient is taking any psychotropic medication or whether or not he/she is hospitalized. The absence of psychotic symptomatology would qualify as a remission only if it lasts for 6 months.

Declaration of Interest

None.

References

Andreasen NC, Carpenter Jr. WT, Kane JM, Lasser RA, Marder SR, Weinberger DR (2005). Remission in schizophrenia: proposed criteria and rationale for consensus. *American Journal of Psychiatry* **162**, 441–449.

APA (1994). *Diagnostic and Statistical Manual of Mental Disorders*, 4th edn. American Psychiatric Association: Washington, DC.

Bloemen OJ, de Koning MB, Schmitz N, Nieman DH, Becker HE, de Haan L, Dingemans P, Linszen DH, van Amelsvoort TA (2010). White-matter markers for psychosis in a prospective ultra-high-risk cohort. *Psychological Medicine* **40**, 1297–1304.

Cuingnet R, Gerardin E, Tessieras J, Auzias G, Lehericy S, Habert MO, Chupin M, Benali H, Colliot O (2011). Automatic classification of patients with Alzheimer's disease from structural MRI: a comparison of ten methods using the ADNI database. *NeuroImage* **56**, 766–781.

Davatzikos C, Fan Y, Wu X, Shen D, Resnick SM (2008). Detection of prodromal Alzheimer's disease via pattern classification of magnetic resonance imaging. *Neurobiology of Aging* **29**, 514–523.

Dazzan P, Lloyd T, Morgan KD, Zanelli J, Morgan C, Orr K, Hutchinson G, Fearon P, Allin M, Rifkin L, McGuire PK, Doody GA, Holloway J, Leff J, Harrison G, Jones PB, Murray RM (2008). Neurological abnormalities and cognitive ability in first-episode psychosis. *British Journal of Psychiatry* **193**, 197–202.

Dazzan P, Morgan KD, Orr KG, Hutchinson G, Chitnis X, Suckling J, Fearon P, Salvo J, McGuire PK, Mallett RM, Jones PB, Leff J, Murray RM (2004). The structural brain correlates of neurological soft signs in AESOP first-episode psychoses study. *Brain* **127**, 143–153.

Eack SM, Pogue-Geile MF, Greenwald DP, Hogarty SS, Keshavan MS (2010). Mechanisms of functional improvement in a 2-year trial of cognitive enhancement therapy for early schizophrenia. *Psychological Medicine*. Published online 22 September 2010. doi:10.1017/S0033291710001765.

Ecker C, Rocha-Rego V, Johnston P, Mourao-Miranda J, Marquand A, Daly EM, Brammer MJ, Murphy C, Murphy DG (2010). Investigating the predictive value of whole-brain structural MR scans in autism: a pattern classification approach. *NeuroImage* **49**, 44–56.

Ellison-Wright I, Glahn DC, Laird AR, Thelen SM, Bullmore E (2008). The anatomy of first-episode and chronic schizophrenia: an anatomical likelihood estimation meta-analysis. *American Journal of Psychiatry* **165**, 1015–1023.

Fan Y, Gur RE, Gur RC, Wu X, Shen D, Calkins ME, Davatzikos C (2008). Unaffected family members and schizophrenia patients share brain structure patterns: a high-dimensional pattern classification study. *Biological Psychiatry* **63**, 118–124.

Harrison G, Hopper K, Craig T, Laska E, Siegel C, Wanderling J, Dube KC, Ganev K, Giel R, an der Heiden W, Holmberg SK, Janca A, Lee PW, Leon CA, Malhotra S, Marsella AJ, Nakane Y, Sartorius N, Shen Y, Skoda C, Thara R, Tsirkin SJ, Varma VK, Walsh D, Wiersma D (2001). Recovery from psychotic illness: a 15- and 25-year international follow-up study. *British Journal of Psychiatry* **178**, 506–517.

Jablensky A, Sartorius N, Ernberg G, Anker M, Korten A, Cooper JE, Day R, Bertelsen A (1992). Schizophrenia: manifestations, incidence and course in different cultures. A World Health Organization ten-country study. *Psychological Medicine. Monograph Supplement* **20**, 1–97.

Karageorgiou E, Schulz SC, Gollub RL, Andreasen NC, Ho B-C, Lauriello J, Calhoun VD, Bockholt HJ, Sponheim SR, Georgopoulos AP (2011). Neuropsychological testing and structural magnetic resonance imaging as diagnostic biomarkers early in the course of schizophrenia and related psychoses. *Neuroinformatics*. Published online 19 January 2011. doi:10.1007/s12021-010-9094-6.

Kasparek T, Thomaz CE, Sato JR, Schwarz D, Janousova E, Marecek R, Prikryl R, Vanicek J, Fujita A, Ceskova E (2011). Maximum-uncertainty linear discrimination analysis of first-episode schizophrenia subjects. *Psychiatry Research: Neuroimaging* **191**, 174–181.

Kloppel S, Stonnington CM, Chu C, Draganski B, Scahill RI, Rohrer JD, Fox NC, Jack Jr. CR, Ashburner J, Frackowiak RS (2008). Automatic classification of MR scans in Alzheimer's disease. *Brain* **131**, 681–689.

Koutsouleris N, Meisenzahl EM, Davatzikos C, Bottlender R, Frodl T, Scheuerecker J, Schmitt G, Zetzsche T, Decker P, Reiser M, Moller HJ, Gaser C (2009). Use of neuroanatomical pattern classification to identify subjects in at-risk mental states of psychosis and predict disease transition. *Archives of General Psychiatry* **66**, 700–712.

LaConte S, Strother S, Cherkassky V, Anderson J, Hu X (2005). Support vector machines for temporal classification of block design fMRI data. *NeuroImage* **26**, 317–329.

Larsen TK, Melle I, Auestad B, Haahr U, Joa I, Johannessen JO, Opjordsmoen S, Rund BR, Rossberg JI,

Simonsen E, Vaglum P, Friis S, McGlashan T (2011). Early detection of psychosis: positive effects on 5-year outcome. *Psychological Medicine* **41**, 1461–1469.

Malla A, Norman R, Bechard-Evans L, Schmitz N, Manchanda R, Cassidy C (2008). Factors influencing relapse during a 2-year follow-up of first-episode psychosis in a specialized early intervention service. *Psychological Medicine* **38**, 1585–1593.

Matthews PM, Honey GD, Bullmore ET (2006). Applications of fMRI in translational medicine and clinical practice. *Nature Reviews Neuroscience* **7**, 732–744.

Mourao-Miranda J, Bokde AL, Born C, Hampel H, Stetter M (2005). Classifying brain states and determining the discriminating activation patterns: Support Vector Machine on functional MRI data. *NeuroImage* **28**, 980–995.

Selten JP, Veen ND, Hoek HW, Laan W, Schols D, van der Tweel I, Feller W, Kahn RS (2007). Early course of schizophrenia in a representative Dutch incidence cohort. *Schizophrenia Research* **97**, 79–87.

Sun D, van Erp TG, Thompson PM, Bearden CE, Daley M, Kushan L, Hardt ME, Nuechterlein KH, Toga AW, Cannon TD (2009). Elucidating a magnetic resonance imaging-based neuroanatomic biomarker for psychosis: classification analysis using probabilistic brain atlas and machine learning algorithms. *Biological Psychiatry* **66**, 1055–1060.

Susser E, Finnerty M, Mojtabai R, Yale S, Conover S, Goetz R, Amador X (2000). Reliability of the life chart schedule for assessment of the long-term course of schizophrenia. *Schizophrenia Research* **42**, 67–77.

van Haren NE, Cahn W, Hulshoff Pol HE, Schnack HG, Caspers E, Lemstra A, Sitskoorn MM, Wiersma D, van den Bosch RJ, Dingemans PM, Schene AH, Kahn RS (2003). Brain volumes as predictor of outcome in recent-onset schizophrenia: a multi-center MRI study. *Schizophrenia Research* **64**, 41–52.

van Haren NE, Hulshoff Pol HE, Schnack HG, Cahn W, Brans R, Carati I, Rais M, Kahn RS (2008). Progressive brain volume loss in schizophrenia over the course of the illness: evidence of maturational abnormalities in early adulthood. *Biological Psychiatry* **63**, 106–113.

Vapnik VN (1995). *The Nature of Statistical Learning Theory*. Springer-Verlag: Berlin.

Vemuri P, Gunter JL, Senjem ML, Whitwell JL, Kantarci K, Knopman DS, Boeve BF, Petersen RC, Jack Jr. CR (2008). Alzheimer's disease diagnosis in individual subjects using structural MR images: validation studies. *NeuroImage* **39**, 1186–1197.

WHO (1992*a*). *The ICD-10 Classification of Mental and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidelines*. World Health Organization: Geneva.

WHO (1992*b*). *Schedules for Clinical Assessment in Neuropsychiatry (SCAN) – Version 2.0*. World Health Organization: Geneva.

Wing JK, Cooper JE, Sartorius N (1974). *The Measurement and Classification of Psychiatric Symptoms*. Cambridge University Press: Cambridge.

Wing JK, Sturt E (1978). *The PSE-ID-CATEGO System: Supplementary Manual*. Medical Research Council: London.